

DATE: 07/11/2002

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/700,843

TIME: 15:39:51

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt

Output Set: N:\CRF3\07112002\I700843.raw

```
3 <110> APPLICANT: LUKACSOVICH, Tamas
              ASZTALOS, Zoltan
      5
              YAMAMOTO, Daisuke
              AWANO, Wakae
      8 <120> TITLE OF INVENTION: A Vector for Gene Trap, and A Method for Gene Trapping by
Using The
              Vector
     11 <130> FILE REFERENCE: 2000-1561A/LC/00653
     13 <140> CURRENT APPLICATION NUMBER: 09/700,843
     14 <141> CURRENT FILING DATE: 2001-02-07
     16 <150> PRIOR APPLICATION NUMBER: PCT/JP99/02683
     17 <151> PRIOR FILING DATÉ: 1999-05-21
     19 <150> PRIOR APPLICATION NUMBER: JP 10-141952
     20 <151> PRIOR FILING DATE: 1998-05-22
     22 <160> NUMBER OF SEQ ID NOS: 7
     24 <170> SOFTWARE: PatentIn Ver. 2.0
     26 <210> SEQ ID NO: 1
     27 <211> LENGTH: 11206
     28 <212> TYPE: DNA
     29 <213> ORGANISM: Artificial sequence
     31 <220> FEATURE:
W--> 32 <221> NAME/KEY: 3'P sequence
     33 <222> LOCATION: (1)..(237)
     35 <220> FEATURE:
W--> 36 <221> NAME/KEY: synthetic splicing acceptor site and stop/start sequence
     37 <222> LOCATION: (238)..(274)
     39 <220> FEATURE:
W--> 40 <221> NAME/KEY: Gal4 gene (coding region and 3'UTR)
     41 <222> LOCATION: (275)..(3164)
     43 <220> FEATURE:
W--> 44 <221> NAME/KEY: hsp70 terminator
     45 <222> LOCATION: (3165)..(3426)
     47 <220> FEATURE:
W--> 48 <221> NAME/KEY: synthetic junction sequence
     49 <222> LOCATION: 3427-3457
     51 <220> FEATURE:
W--> 52 <221> NAME/KEY: heat shock promoter directed neomycine resistance gene on complementer
     53
              strand
     54 <222> LOCATION: (3458)..(4907)
     56 <220> FEATURE:
W--> 57 <221> NAME/KEY: mini-white gene
     58 <222> LOCATION: (4908)..(8275)
     60 <220> FEATURE:
```

W--> 61 <221> NAME/KEY: synthetic splicing donor site

RAW SEQUENCE LISTING

DATE: 07/11/2002 TIME: 15:39:51

PATENT APPLICATION: US/09/700,843

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt

Output Set: N:\CRF3\07112002\I700843.raw

```
62 <222> LOCATION: (8276)..(8299)
     64 <220> FEATURE:
W--> 65 <221> NAME/KEY: 5'P sequence
     66 <222> LOCATION: (8300)..(8446)
     68 <220> FEATURE:
W--> 69 <221> NAME/KEY: bacterial part of pCasper3 shuttle vector including complete pUC8
W--> 70
              sequence
    71 <222> LOCATION: (8447)..(11206)
    73 <220> FEATURE:
W--> 74 <221> NAME/KEY: synthetic DNA
     75 <222> LOCATION: (238)..(274)
              (3427)..(3457)
              (4908)..(4914)(8276)..(8299)
     77
     78 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
     80 <400> SEQUENCE: 1
     81 catgatgaaa taacataagg tggtcccgtc ggcaagagac atccacttaa cgtatgcttg
                                                                                60
     82 caataagtgc gagtgaaagg aatagtattc tgagtgtcgt attgagtctg agtgagacag
                                                                               120
     83 cgatatgatt gttgattaac ccttagcatg tccgtggggt ttgaattaac tcataatatt
                                                                               180
     84 aattagacga aattatttt aaagttttat ttttaataat ttgcgagtac gcaaagctct
                                                                               240
     85 ttctcttaca ggtcgaattg atgtgatgga tccaatgaag ctactgtctt ctatcgaaca
                                                                               300
                                                                               360
     86 agcatgcgat atttgccgac ttaaaaagct caagtgctcc aaagaaaaac cgaagtgcgc
     87 caagtgtctg aagaacaact gggagtgtcg ctactctccc aaaaccaaaa ggtctccgct
                                                                               420
                                                                               480
     88 qactaqqqca catctgacag aagtggaatc aaggctagaa agactggaac agctatttct
     89 actgattttt cctcgagaag accttgacat gattttgaaa atggattctt tacaggatat
                                                                               540
     90 aaaagcattg ttaacaggat tatttgtaca agataatgtg aataaagatg ccgtcacaga
                                                                               600
     91 tagattggct tcagtggaga ctgatatgcc tctaacattg agacagcata gaataagtgc
                                                                               660
     92 gacatcatca toggaagaga gtagtaacaa aggtcaaaga cagttgactg tatogattga
                                                                               720
     93 ctcggcagct catcatgata actccacaat tccgttggat tttatgccca qqqatqctct
                                                                               780
                                                                               840
     94 tcatggattt gattggtctg aagaggatga catgtcggat ggcttgccct tcctgaaaac
     95 ggaccccaac aataatgggt tetttggega eggttetete ttatgtatte ttegatetat
                                                                               900
     96 tggctttaaa ccggaaaatt acacgaactc taacgttaac aggctcccga ccatgattac
                                                                               960
     97 ggatagatac acgttggctt ctagatccac aacatcccgt ttacttcaaa gttatctcaa
                                                                              1020
     98 taattttcac ccctactgcc ctatcgtgca ctcaccgacg ctaatgatgt tgtataataa
                                                                              1080
     99 ccagattgaa atcgcgtcga aggatcaatg gcaaatcctt tttaactgca tattagccat
                                                                              1140
     100 tggagcctgg tgtatagagg gggaatctac tgatatagat gttttttact atcaaaatgc
                                                                               1200
                                                                               1260
     101 taaatctcat ttgacgagca aggtcttcga gtcaggttcc ataattttgg tgacagccct
                                                                               1320
     102 acatettetg tegegatata caeagtggag geagaaaaca aatactaget ataattttea
     103 cagettttee ataagaatgg ceatateatt gggettgaat agggaeetee eetegteett
                                                                               1380
     104 cagtgatage ageattetgg aacaaagaeg cegaatttgg tggtetgtet actettggga
                                                                               1440
     105 gatecaattg teeetgettt atggtegate catecagett teteagaata caateteett
                                                                               1500
     106 cccttcttct gtcgacgatg tgcagcgtac cacaacaggt cccaccatat atcatggcat
                                                                               1560
     107 cattgaaaca gcaaggctct tacaagtttt cacaaaaatc tatgaactag acaaaacagt
                                                                               1620
     108 aactgcagaa aaaagtccta tatgtgcaaa aaaatgcttg atgatttgta atgagattga
                                                                               1680
     109 ggaggtttcg agacaggcac caaagttttt acaaatggat atttccacca ccgctctaac
                                                                               1740
     110 caatttgttg aaggaacacc cttggctatc ctttacaaga ttcgaactga agtggaaaca
                                                                               1800
     111 gttgtctctt atcatttatg tattaagaga ttttttcact aattttaccc agaaaaagtc
                                                                               1860
     112 acaactaqaa caqqatcaaa atqatcatca aagttatgaa gttaaacgat gctccatcat
                                                                               1920
     113 gttaagcgat gcagcacaaa gaactgttat gtctgtaagt agctatatgg acaatcataa
                                                                               1980
                                                                               2040
     114 tgtcacccca tattttgcct ggaattgttc ttattacttg ttcaatgcag tcctagtacc
```



PATENT APPLICATION: US/09/700,843

DATE: 07/11/2002 TIME: 15:39:51

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt
Output Set: N:\CRF3\07112002\I700843.raw

115	cataaagact	ctactctcaa	actcaaaatc	gaatgctgag	aataacgaga	ccgcacaatt	2100
116	attacaacaa	attaacactg	ttctgatgct	attaaaaaaa	ctggccactt	ttaaaatcca	2160
117	gacttgtgaa	aaatacattc	aagtactgga	agaggtatgt	gcgccgtttc	tgttatcaca	2220
118	gtgtgcaatc	ccattaccgc	atatcagtta	taacaatagt	aatggtagcg	ccattaaaaa	2280
119	tattgtcggt	tctgcaacta	tcgcccaata	ccctactctt	ccggaggaaa	atgtcaacaa	2340
120	tatcagtgtt	aaatatgttt	ctcctggctc	agtagggcct	tcacctgtgc	cattgaaatc	2400
121	aggagcaagt	ttcagtgatc	tagtcaagct	gttatctaac	cgtccaccct	ctcgtaactc	2460
122	tccagtgaca	ataccaagaa	gcacaccttc	gcatcgctca	gtcacgcctt	ttctagggca	2520
123	acagcaacag	ctgcaatcat	tagtgccact	gaccccgtct	gctttgtttg	gtggcgccaa	2580
124	ttttaatcaa	agtgggaata	ttgctgatag	ctcattgtcc	ttcactttca	ctaacagtag	2640
125	caacggtccg	aacctcataa	caactcaaac	aaattctcaa	gcgctttcac	aaccaattgc	2700
126	ctcctctaac	gttcatgata	acttcatgaa	taatgaaatc	acggctagta	aaattgatga	2760
127	tggtaataat	tcaaaaccac	tgtcacctgg	ttggacggac	caaactgcgt	ataacgcgtt	2820
				tacaatggat			2880
129	cgatgatgaa	gataccccac	caaacccaaa	aaaagagtaa	aatgaatcgt	agatactgaa	2940
130	aaaccccgca	agttcacttc	aactgtgcat	cgtgcaccat	ctcaatttct	ttcatttata	3000
131	catcgttttg	ccttctttta	tgtaactata	ctcctctaag	tttcaatctt	ggccatgtaa	3060
132	cctctgatct	atagaatttt	ttaaatgact	agaattaatg	cccatctttt	ttttggacct	3120
133	aaattcttca	tgaaaatata	ttacgagggc	ttattcagaa	gcttatcgat	accgtcgact	3180
134	aaagccaaat	agaaattatt	cagttctggc	ttaagttttt	aaaagtgata	ttatttattt	3240
135	ggttgtaacc	aaccaaaaga	atgtaaataa	ctaatacata	attatgttag	ttttaagtta	3300
136	gcaacaaatt	gattttagct	atattagcta	cttggttaat	aaatagaata	tatttattta	3360
137	aagataattc	gtttttattg	tcagggagtg	agtttgctta	aaaactcgtt	tagatccact	3420
				aaggagggcg			3480
139	cccgcgctgg	aggatcatcc	agccggcgtc	ccggaaaacg	attccgaagc	ccaacctttc	3540
140	atagaaggcg	gcggtggaat	cgaaatctcg	tgatggcagg	ttgggcgtcg	cttggtcggt	3600
141	catttcgaac	cccagagtcc	cgctcagaag	aactcgtcaa	gaaggcgata	gaaggcgatg	3660
142	cgctgcgaat	cgggagcggc	gataccgtaa	agcacgagga	agcggtcagc	ccattcgccg	3720
143	ccaagctctt	cagcaatatc	acgggtagcc	aacgctatgt	cctgatagcg	gtccgccaca	3780
144	cccaqccggc	cacagtcgat	gaatccagaa	aagcggccat	tttccaccat	gatattcggc	3840
145	aagcaggcat	cgccatgggt	cacgacgaga	tectegeegt	cgggcatgcg	cgccttgagc	3900
146	ctggcgaaca	gttcggctgg	cgcgagcccc	tgatgctctt	cgtccagatc	atcctgatcg	3960
147	acaagaccgg	cttccatccg	agtacgtgct	cgctcgatgc	gatgtttcgc	ttggtggtcg	4020
148	aatgggcagg	tagccggatc	aagcgtatgc	agccgccgca	ttgcatcagc	catgatggat	4080
149	actttctcgg	caggagcaag	gtgagatgac	aggagatcct	gccccggcac	ttcgcccaat	4140
150	agcagccagt	cccttcccgc	ttcagtgaca	acgtcgagca	cagctgcgca	aggaacgccc	4200
151	gtcgtggcca	gccacgatag	ccgcgctgcc	togtootgca	gttcattcag	ggcaccggac	4260
152	aggtcggtct	tgacaaaaag	aaccgggcgc	ccctgcgctg	acagccggaa	cacggcggca	4320
153	tcagagcagc	cgattgtctg	ttgtgcccag	tcatagccga	atagcctctc	cacccaagcg	4380
154	gccggagaac	ctgcgtgcaa	tccatcttgt	tcaatcatgc	gaaacgatcc	tcatcctgtc	4440
155	tcttgatcag	atcccctatt	cagagttctc	ttcttgtatt	caataattac	ttcttggcag	4500
				tgctggttac			4560
157	cttgcacttt	actgcagatt	gtttagcttg	ttcagctgcg	cttgtttatt	tgcttagctt	4620
158	tcgcttagcg	acgtgttcac	tttgcttgtt	tgaattgaat	tgtcgctccg	tagacgaagc	4680
159	gcctctattt	atactccggc	gctcttttcg	cgaacattcg	aggcgcgctc	tctcgaacca	4740
160	acgagagcag	tatgccgttt	actgtgtgac	agagtgagag	agcattagtg	cagagaggga	4800
161	gagacccaaa	aagaaaagag	agaataacga	ataacggcca	gagaaatttc	tcgagttttc	4860
162	tttctgccaa	acaaatgacc	taccacaata	accagtttgt	tttgggatct	agtccctaat	4920
163	tctagtatgt	atgtaagtta	ataaaaccct	tttttggaga	atgtagattt	aaaaaaacat	4980

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/700,843

DATE: 07/11/2002
TIME: 15:39:51

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt
Output Set: N:\CRF3\07112002\I700843.raw

			tgcactggac				5040
			ttgaagtacc				5100
166	gacattccac	tcaaagtcag	cgctgtttgc	ctccttctct	gtccacagaa	atatcgccgt	5160
			ctatctcttt				5220
			tttgtcagcg				5280
			gctgtgccaa				5340
			gattggtttt				5400
			gatcaggagc				5460
			gtgagtagta				5520
			catcaaattg				5580
174	atcccaaacc	gcaatcgcac	aaacaaatag	tgacacgaaa	cagattattc	tggtagctgt	5640
175	gctcgctata	taagacaatt	tttaagatca	tatcatgatc	aagacatcta	aaggcattca	5700
176	ttttcgacta	cattctttt	tacaaaaaat	ataacaacca	gatattttaa	gctgatccta	5760
177	gatgcacaaa	aaataaataa	aagtataaac	ctacttcgta	ggatacttcg	ttttgttcgg	5820
178	ggttagatga	gcataacgct	tgtagttgat	atttgagatc	ccctatcatt	gcagggtgac	5880
			gcattaacca				5940
			cggaggactc				6000
181	cctcacctat	gcctggcaca	atatggacat	ctttggggcg	gtcaatcagc	cgggctccgg	6060
182	atggcggcag	ctggtcaacc	ggacacgcgg	actattctgc	aacgagcgac	acataccggc	6120
183	gcccaggaaa	catttgctca	agaacggtga	gtttctattc	gcagtcggct	gatctgtgtg	6180
			attaccaatt				6240
			gggcagttcc				6300
			gggcatccaa				6360
			ggagatgcag				6420
188	ctctttatcg	gctccctaac	ggccagggaa	cacctgattt	tccaggccat	ggtgcggatg	6480
			gcagcgagtg				6540
190	tcgctcagca	aatgtcagca	cacgatcatc	ggtgtgcccg	gcagggtgaa	aggtctgtcc	6600
			ggcattcgcc				6660
			cggactggac				6720
			gggcaagacc				6780
194	gagctgtttg	agctctttga	caagatcctt	ctgatggccg	agggcagggt	agctttcttg	6840
195	ggcactccca	gcgaagccgt	cgacttcttt	tcctagtgag	ttcgatgtgt	ttattaaggg	6900
196	tatctagcat	tacattacat	ctcaactcct	atccagcgtg	ggtgcccagt	gtcctaccaa	6960
197	ctacaatccg	gcggactttt	acgtacaggt	gttggccgtt	gtgcccggac	gggagatcga	7020
			agatatgcga	_	-		7080
			ccaaaaattt				7140
200	gtacacctac	aaggccacct	ggttcatgca	gttccgggcg	gtcctgtggc	gatcctggct	7200
			tcctcgtaaa				7260
202	ttccagtgga	aacaaatgat	ataacgctta	caattcttgg	aaacaaattc	gctagatttt	7320
			acacccttct				7380
			aatttcattt				7440
			tttaatttgc				7500
			caagtgggcg				7560
			caaaacgtct				7620
			tttactaact				7680
			tatgagggag				7740
			tgccgaatta				7800
			gatcggactg				7860
212	ctggcgctgg	tcactctggt	ggccaatgtg	tcaacgtcct	toggatatot	aatatcctgc	7920

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/700,843

DATE: 07/11/2002 TIME: 15:39:51

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt
Output Set: N:\CRF3\07112002\I700843.raw

213	gccagctcct	cgacctcgat	ggcgctgtct	gtgggtccgc	cggttatcat	accattcctg	7980
				tcggtgccag			8040
				ggtctgctga			8100
				aacaccacgt			8160
				gatctgccgc			8220
218	attctcatcg	tgagcttccg	ggtgctcgca	tatctggctc	taagacttcg	ggcccgacgc	8280
				taagggttaa			8340
				aacgtgcact			8400
				atgttatttc			8460
				taagcgtctc			8520
				taatttcgcc			8580
				ggaaccatga			8640
				agaaggcgtt			8700
				gtcagaggtg			8760
227	aaagatacca	ggcgtttccc	cctggaagct	ccctcgtgcg	ctctcctgtt	ccgaccctgc	8820
228	cgcttaccgg	atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	tctcaatgct	8880
229	cacgctgtag	gtatctcagt	tcggtgtagg	tcgttcgctc	caagctgggc	tgtgtgcacg	8940
				tatccggtaa			9000
231	cggtaagaca	cgacttatcg	ccactggcag	cagccactgg	taacaggatt	agcagagcga	9060
232	ggtatgtagg	cggtgctaca	gagttcttga	agtggtggcc	taactacggc	tacactagaa	9120
233	ggacagtatt	tggtatctgc	gctctgctga	agccagttac	cttcggaaaa	agagttggta	9180
234	gctcttgatc	cggcaaacaa	accaccgctg	gtagcggtgg	tttttttgtt	tgcaagcagc	9240
				aagatccttt			9300
236	acgctcagtg	gaacgaaaac	tcacgttaag	ggattttggt	catgagatta	tcaaaaagga	9360
				gaagttttaa			9420
				taatcagtga			9480
239	gtctatttcg	ttcatccata	gttgcctgac	tccccgtcgt	gtagataact	acgatacggg	9540
				tgataccgcg			9600
				gaagggccga			9660
242	ctttatccgc	ctccatccag	tctattaatt	gttgccggga	agctagagta	agtagttcgc	9720
				ttgctacagg			9780
				cccaacgatc			9840
				taggtactac			9900
				cagcactgca			9960
				agtactcaac			10020
248	gtatgcggcg	accgagttgc	tcttgcccgg	cgtcaacacg	ggataatacc	gcgccacata	10080
249	gcagaacttt	aaaagtgctc	atcattggaa	aacgttcttc	ggggcgaaaa	ctctcaagga	10140
				aacccactcg			10200
				gagcaaaaac			10260
				gaatactcat			10320
253	attgaagcat	ttatcagggt	tattgtctca	tgagcggata	catatttgaa	tgtatttaga	10380
				ttccccgaaa			10440
255	aaaccattat	tatcatgaca	ttaacctata	aaaataggcg	tatcacgagg	ccctttcgtc	10500
				tctgacacat			10560
				gacaagcccg			10620
				cggcatcaga			10680
				tegegeggaa			10740
260	gtcgtcgaac	aaaaggtgaa	tgtgttgcgg	agagcgggtg	ggagacagcg	aaagagcaac	10800
261	tacgaaacgt	ggtgtggtgg	aggtgaatta	tgaagagggc	gcgcgatttg	aaaagtatgt	10860

VERIFICATION SUMMARY

DATE: 07/11/2002

PATENT APPLICATION: US/09/700,843

TIME: 15:39:52

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt

Output Set: N:\CRF3\07112002\1700843.raw

```
L:32 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:36 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:40 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:44 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:48 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:50 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:57 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:61 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:65 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:69 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:70 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:1
L:74 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
```